PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau
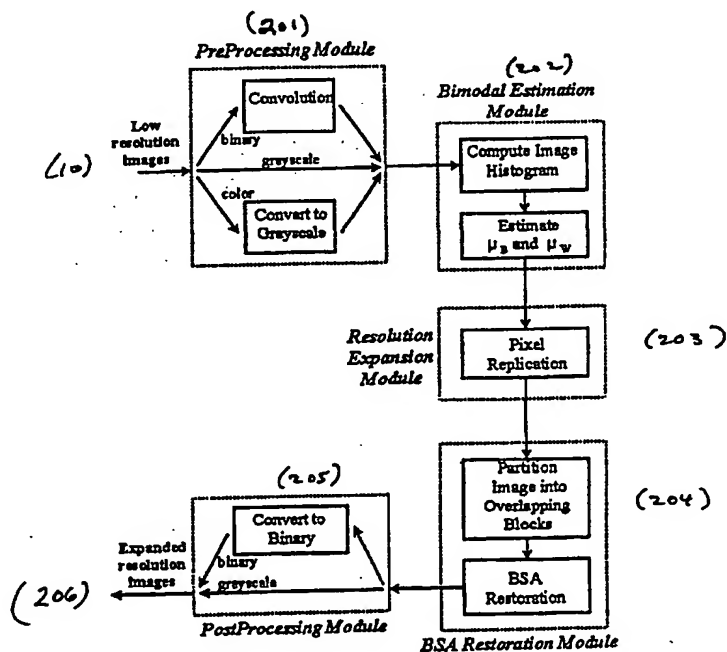
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

/

(54) Title: METHOD AND APPARATUS FOR RESTORATION OF LOW RESOLUTION IMAGES

(57) Abstract

The definition of a low–resolution image (10) may be improved by taking advantage of characteristics of its bimodal distribution (202) so as to produce an expanded image (206) with improved definition. This can not only reconstruct image information deteriorated by low–resolution devices, but also can refine original low spatial resolution so as to significantly improve image quality. An image is improved by iteratively solving a nonlinear optimization problem using a preferred Bimodal–Smooth–Average scoring method. Among the applications for the technique are improved optical character recognition, restoration of binary images, and restoration of video frames.

Block diagram of the text resolution expansion system.

TITLE: METHOD AND APPARATUS FOR RESTORATION OF LOW RESOLUTION IMAGES

INVENTORS: Paul THOUIN and Chein-I CHANG

This application is entitled to the benefit of the filing date of U.S. Provisional Application Serial No. 60/127,316, filed April 1, 1999.

FIELD AND BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates in general to image processing, and more particularly to creating a higher resolution image from a lower resolution image.

Background Information

The invention described and claimed herein comprises a novel method and apparatus for improving the definition of a low-resolution image by taking advantage of characteristics of its bimodal distribution so as to produce an expanded image with improved definition. This can not only reconstruct image information deteriorated by low-resolution devices, but also can refine original low spatial resolution images so as to

1

significantly improve image quality. An image is improved by iteratively solving a nonlinear optimization problem using a preferred Bimodal-Smooth-Average scoring method. Among the applications for the technique are improved optical character recognition, restoration of binary images and video frames.

As used herein, the terms "low" and "high" are comparative only: they refer to the comparison between two images and are not meant to imply any absolute degree of resolution.

As described in greater detail below, image acquisition consists of converting a continuous image into discrete values obtained from a group of sensor elements. While generally cheaper to produce and transmit, a low resolution imaging system generally produces a less accurate representation of the continuous image than would a high resolution imaging system. Common methods of expanding low resolution images to high resolution images by interpolation typically smooth over important details in the process. Linear interpolation tends to smooth image data at transition regions, resulting in blurry images. Cubic spline expansion allows for sharp transitions, but tends to produce a ringing effect at discontinuities.

Many applications (including OCR, video surveillance and video compression algorithms) would benefit from improved expansion techniques.

2

Processing of document images has traditionally been performed using this

binary format due to the savings in disk storage and computer processing. As com-

puter performance and storage capacity have dramatically increased recently, grayscale

scanning and processing of document images have become more prevalent. Text images

traditionally have been obtained from documents. However due to the current increase

in digital video a significant amount of text is found in this type of imagery as well.

Document images are typically acquired by a scanner, whereas video frames are most

often captured by a digital camera. The image acquisition process in both cases consists

of converting a continuous image into discrete values obtained from a group of sensor

elements. Each sensor element produces a value which is a function of the amount

of light incident on the device. For 8-bit grayscale quantization, the allowable range

of values for each sensor are integers from 0 (black) to 255 (white). The sensors are

typically arranged in a non-overlapping grid of square elements, smaller elements result

in higher resolution imagery.

Text image resolution expansion has become increasingly important in a number of areas of image processing. Optical Character Recognition (OCR) of document images continues to be of great importance as we attempt to become a paperless society. Restoring text from video surveillance imagery is often crucial to law enforcement agencies. Digital video compression algorithms can also benefit from successful text resolution expansion techniques. Common methods of interpolation, which were not designed specifically for text images, typically smooth over the important details and produce inadequate expansion. This paper proposes a new nonlinear restoration technique for text images, which creates smooth foreground and background regions while preserving sharp edge transitions.

There has been significant research in both the text enhancement and resolution expansion fields. Some text enhancement efforts focus on fixing broken or touching characters [1] [2]. Other methods [3] [4] use degraded samples to model the image and then solve for the restored image using using a recursive technique. To remove the effects of noise, [5] uses a morphological filter and [6] uses pixel patterns to improve OCR results. A variety of methods have been proposed in order to improve contrast within text images including quadratic filters [7], soft morphological filters [8], non-linear mapping [9], and using a multi-resolution pyramid approach and fuzzy edge detectors [10]. Numerous resolution expansion methods have been published in the literature as well [11]-[17]. Linear interpolation tends to smooth the image data at transition regions

and results in a high-resolution image that appears blurry. Cubic spline expansion allows for sharp transitions, but tends to produce a ringing effect at these discontinuities.

A number of research efforts investigated combining text enhancement with resolution expansion in order to improve low-resolution text images. Shannon interpolation is performed with text separation from the image background in [18] to improve the OCR accuracy of digital video. Deblurring is combined with linear interpolation in [19] to enhance document images obtained from digital cameras. Both of these methods essentially perform the text enhancement and resolution expansion as two separate steps. The proposed method computes resolution expansion using an algorithm specifically designed to enhance text images to perform both of these steps simultaneously.

The goal of resolution expansion is to create an expanded image with improved definition from observed low-resolution imagery. Acquisition of this low-resolution imagery can be modeled by averaging a block of pixels within a high-resolution image. Resolution expansion is an ill-posed inverse problem. For a given low-resolution image, a virtually infinite set of expanded images can be generated by the observed data.

SUMMARY OF THE INVENTION

The foregoing problems are overcome, and other advantages are provided by a method and apparatus for enhancing an image by iteratively solving a nonlinear optimization problem using a preferred Bimodal-Smooth-Average (BSA) scoring method, in accordance with the invention, which minimizes a BSA score, defined as the weighted sum of separate bimodal, smoothness and average measures. This can not only reconstruct image information deteriorated by low-resolution devices, but also can refine original low spatial resolution so as to significantly improve image quality.

It is an object of the invention to restore a high-resolution image given a low-resolution image.

It is an object of the invention to provide a method and apparatus for improving the definition of a low-resolution image so as to produce an image with improved definition.

It is a further object of the invention to provide a method and apparatus for reconstructing image information deteriorated by low-resolution devices.

It is a further object of the invention to provide a method and apparatus for refining original low spatial resolution so as to significantly improve image quality.

6

It is a further object of the invention to automatically enhance text within document or video images so as to improve the accuracy of subsequent optical character recognition.

It is a further object of the invention to provide a method and apparatus for improving restoration of video images.

A principal feature of the invention is the iterative solution of a nonlinear optimization problem using a preferred Bimodal-Smooth-Average scoring method.

Among the advantages of the invention are automated enhancement of images.

These and other objects, features and advantages which will be apparent from the discussion which follows are achieved, in accordance with the invention, by providing a novel method and apparatus for improving image quality by iteratively solving a nonlinear optimization problem, using a Bimodal-Smooth-Average scoring method.

The various features of novelty which characterize the invention are pointed out with particularity in the claims annexed to and forming a part of this disclosure. For a better understanding of the invention, its advantages and objects, reference is made to the accompanying drawings and descriptive matter in which a

7

preferred embodiment of the invention is illustrated.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and still other objects of this invention will become apparent, along with various advantages and features of novelty residing in the present embodiments, from study of the following drawings, in which:

Figure 1 is a schematic of a system suitable for carrying out the invention.

Figure 2 is a block diagram of the text resolution expansion system of the invention.

Figure 3 is a flow chart of the BSA process of the invention.

Figure 4 is an illustration of the results of a BSA restoration experiment, showing the progressive improvement of resolution with iterations.

Figures 5 illustrates the performance of the BSA process as compared to prior art techniques of linear interpolation and cubic spline in grayscale document restoration and grayscale video frame restoration, respectively.

Figures 6 and 7 illustrate the performance of optical character recognition following cubic spline (Figure 6) versus BSA (Figure 7) image enhancement.

Figure 8 compares OCR accuracy of spline vs. BSA for gray-level images and binary document images.

Figure 9 illustrates experimental results of binary document restoration.

Figures 10 and 11 illustrate experimental results of restoration of video frames.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The invention is a novel method and apparatus which improves the definition of a low-resolution image by taking advantage of characteristics of its bimodal distribution so as to produce an expanded image with improved definition. This can not only reconstruct image information deteriorated by low-resolution devices, but also can refine original low spatial resolution so as to significantly improve image quality.

Referring to Figure 1, the invention may be illustrated in the context of an image acquisition system (1) providing image data (10) which serves as an input to a general purpose computer (100) programmed to carry out the following processes.

Image acquisition system (1) comprises a plurality of image sensors (2) which generate image data (10) concerning the image. This image data (10) serves as input to computer (100) via input means (101). Computer (100) includes processing means (102), programmed to execute the steps set forth herein so as to generate a high resolution image, which may be displayed or otherwise used via output means (103).

Referring to Figure 2, a text resolution expansion system receives low-resolution image (10), preprocesses the image in preprocessing module (201) and outputs the result to bimodal

10

estimation module (202); the bimodal estimation module (202) processes the data and outputs the result to resolution expansion module (203) for pixel replication and output to BSA restoration module (204) where the image is partitioned into overlapping blocks prior to BSA restoration; the BSA restoration module then outputs the result to a post-processing module (205) which provides an expanded resolution image (206) for display or other use.

The operation of the BSA restoration module (203) is illustrated in greater detail in Figure 3. An initial image (301) is provided and an initial BSA score is computed (302); an image update is then computed (303) so as to produce an updated image (304). A new BSA score is computed for the updated image (305) and compared to the prior BSA score (306); if the new score does not represent an improvement over the prior score by a predetermined amount, the restoration process is complete; otherwise, another image update is computed (303) and the process continues iteratively until the improvement is below the predetermined amount.

The image acquisition process consists of converting a continuous image into discrete values obtained from a group of sensor elements. Each sensor element produces a value which is a function of the amount of light incident on the device. For 8-bit grayscale quantization, the allowable range of values for each sensor are integers from 0 (black) to 255 (white). The sensors are typically arranged in a non-overlapping grid of square

elements, smaller elements result in higher resolution imagery. ~~In a~~ high-resolution imaging system ~~~~ the number of sensors is adequate to represent the desired text image. The majority of pixels within the image are either white or black, with a small number of gray pixels occurring at the edges. ~~In a~~ low-resolution imaging system ~~~~ the number of sensors has been reduced ~~~~ ~~~~. This low-resolution acquisition results in significant blockiness and is insufficient to accurately represent this image. Each sensor element effectively averages the image within its section of the grid, resulting in an increased amount of gray pixels. Low-resolution imaging can therefore be thought of as block-averaging high-resolution images.

The problem ~~~~ is to restore the high-resolution image $HI_{qr,qc}$

given only the low-resolution image $LI_{r,c}$, where r and c are the number of rows and columns in the low-resolution image and q is the resolution expansion factor. The image acquisition process of obtaining $LI_{r,c}$ from $HI_{qr,qc}$ is given by

$$LI_{r,c} = \frac{1}{q^2} \sum_{s=q \cdot r}^{(qr+q-1)} \sum_{t=q \cdot c}^{(qc+q-1)} HI_{s,t} \qquad (1)$$

The value of $LI_{r,c}$ is the average of the high-resolution pixels within the $q \times q$ neighborhood. Eq. (1) represents a typical image restoration problem where we are required to restore the $HI_{qr,qc}$ based on the observed $LI_{r,c}$ via the relationship described by this equation. Since there are a great number of high-resolution images which may satisfy the constraint of the observed low-resolution image given by Eq. (1), image restoration is generally an ill-posed inverse problem.

The system presented in ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬

▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬

▬▬▬block diagram▬▬▬ in Fig. ▬ comprises five modules, the PreProcessing Module, the Bimodal Estimation Module, the Resolution Expansion Module, the BSA Restoration Module, and the PostProcessing Module. A low-resolution image is initially input to the PreProcessing Module where binary and color images are converted to grayscale images. The image histogram is computed by the Bimodal Estimation Module to estimate the means of the black and white pixel distributions. The initial image expansion is performed by the Resolution Expansion Module which uses pixel replication to create a high-resolution image from the low-resolution original. The image is divided into overlapping blocks of pixels which are restored independently by the BSA algorithm in the BSA Restoration Module. When a binary expanded image is desired, the PostProcessing Module performs a grayscale-to-binary conversion. The next section details the image expansion system for grayscale images, followed by a section on binary image expansion.

3   Grayscale Image Resolution Expansion

The image resolution expansion process for grayscale images, which includes both video frames and grayscale document images, is described in this section. It should be noted that our system depicted in Fig. 7 processes grayscale and color images in a very similar fashion. The only difference is in the PreProcessing Module where color

images are initially converted to grayscale images using standard techniques. These formerly color images are processed throughout the system as grayscale and the resulting high-resolution expanded images are grayscale as well. Since the goal of our system is to improve automated OCR results, this loss of color information is not a concern. Once the system has a grayscale image, its histogram is then computed by the Bimodal Estimation Module. The distribution of a text image is typically bimodal with a large white peak corresponding to the background and a smaller dark peak corresponding to the text pixels. The peak of the white distribution μ_W and the peak of the black distribution μ_B are estimated from the calculated histogram. These values are required later by the system. Pixel replication, where the value of each high-resolution pixel is equal to its corresponding low-resolution pixel, is performed in the Resolution Expansion Module to form the initial expanded image. The BSA Restoration Module partitions the image into overlapping blocks and restores the blocks independently. Finally, the PostProcessing Module converts the high-resolution grayscale image to a binary image if a binary expanded image is desirable. In what follows, the BSA Restoration Module will be described in detail.

A block diagram of the BSA Restoration process is shown in Fig. 3 where I_i represents the image at iteration i, BSA_i is the score at iteration i, and δ_i is the change to the image at iteration i. The restoration process can be summarized as follows. The initial score BSA_0 is computed from the original image I_0. At each iteration, the image update δ_i and new score BSA_i are computed. The iterations continue to minimize the

BSA score until a convergence is reached. At this stage, a restored image is produced.

The scoring function used by the BSA Restoration process is designed to measure how well a group of pixels within an image represent the desired properties of a text image. This function, referred to as the BSA scoring function, is expressed as the weighted sum of a Bimodal score B, a Smoothness score S, and an Average score A. More precisely, the BSA score is defined by

$$BSA(x) = \lambda_1 B(x) + \lambda_2 S(x) + \lambda_3 A(x) \qquad (2)$$

where λ_1, λ_2, and λ_3 are Lagrange multipliers and x is a block of pixels to be restored. Throughout this paper, a block x is defined both as a group of 4×4 low-resolution pixels or as the $4q \times 4q$ high-resolution pixels that are derived from them. The 4×4 size was specifically chosen because it contains enough pixels to adequately measure text characteristics but is not too large to be computationally burdensome. The goal of the restoration process is to iteratively solve for the block of pixels x that minimizes the $BSA(x)$ score given by Eq. (2). Each of the three scores will be briefly discussed in the following subsections followed by a description of an iterative minimization procedure.

A. The Bimodal Score

The typical distribution of a text image contains two peaks, a large one at μ_{white}, which normally represent the page's background, and a secondary peak at μ_{black} representing the foreground text █████████████ From the histogram of the given low-resolution text image, estimates of the means for the black and white distributions are calculated. These means are used to compute the bimodal score $B(x)$, which measures how far an image block x is from bimodal. The bimodal score used in this paper is defined by

$$B(x) = \sum_{r,c}(x_{r,c} - \mu_{black})^2(x_{r,c} - \mu_{white})^2 \tag{3}$$

where r and c are the row and column indices within the block being evaluated.

The bimodal function for $\mu_{black} = 50$ and $\mu_{white} = 255$ is plotted in Fig. 4. When a pixel value within x is close to either μ_{black} or μ_{white}, its contribution to $B(x)$ is minimal. The bimodal minimum score of $B(x) = 0$ means that the image is perfectly bimodal, the value of every pixel is equal to either μ_{white} or μ_{black}. Solving for the block of pixels x that minimizes $B(x)$ produces a strongly bimodal image, which is one of the desired properties of this proposed text restoration technique.

To minimize $B(x)$, both the first and second derivatives of the bimodal score will be computed. The score given by Eq. (3) is differentiable, the partial derivative of $B(x)$

with respect to pixel $x_{r,c}$ is

$$\frac{\partial B(x)}{\partial x_{r,c}} = 4x_{r,c}^3 - 6(\mu_{white} + \mu_{black})x_{r,c}^2 \qquad (4)$$
$$+2(\mu_{white}^2 + 4\mu_{white}\mu_{black} + \mu_{black}^2)x_{r,c} - 2\mu_{white}\mu_{black}(\mu_{white} + \mu_{black})$$

and the second partial derivative for the bimodal score is

$$\frac{\partial^2 B(x)}{\partial x_{r,c}^2} = 12x_{r,c}^2 - 12(\mu_{white} + \mu_{black})x_{r,c} + 2(\mu_{white}^2 + 4\mu_{white}\mu_{black} + \mu_{black}^2) \qquad (5)$$

Partial derivatives of the bimodal score $B(x)$ are independent of neighboring pixels. The estimated means of the bimodal distribution, μ_{black} and μ_{white}, are known a priori and their appropriate constants can be pre-computed.

B. The Smoothness Score

With the exception of edges, text images tend to be very smooth in both the foreground and background regions which results in neighbors with similar values. A smoothness score, which is computed for each block of pixels, is introduced to measure this feature. For this proposed algorithm, a simple statistic using only the four nearest neighbors of each pixel is used. Other more sophisticated smoothness measures could be implemented as well. The smoothness score $S(x)$ used by this technique is given by

$$S(x) = \sum_{r,c}[(x_{r-1,c} - x_{r,c})^2 + (x_{r,c-1} - x_{r,c})^2 + (x_{r,c+1} - x_{r,c})^2 + (x_{r+1,c} - x_{r,c})^2] \qquad (6)$$

where r and c are the row and column indices within the block being evaluated. The minimum value of $S(x) = 0$ occurs when all pixels have identical values.

The first and second partial derivatives are calculated in a straightforward manner.

The first partial derivative of this smoothing score with respect to pixel $x_{r,c}$ is

$$\frac{\partial S(x)}{\partial x_{r,c}} = -2(x_{r-1,c} + x_{r,c-1} + x_{r,c+1} + x_{r+1,c}) + 8x_{r,c} \tag{7}$$

The second partial derivatives are nonzero only when

$$\frac{\partial^2 S(x)}{\partial x_{r,c}^2} = 8 \qquad \frac{\partial^2 S(x)}{\partial x_{r,c} \partial x_{r\pm1,c}} = -2 \qquad \frac{\partial^2 S(x)}{\partial x_{r,c} \partial x_{r,c\pm1}} = -2 \tag{8}$$

and are equal to zero for all other combinations. The first partial derivative of the smoothness score $S(x)$ is a function of the four neighboring pixels. The second partial derivatives are non-zero only with respect to their corresponding neighbors.

C. The Average Constraint Score

It is reasonable to require that the average of a group of high-resolution pixels is close to the original value of the low-resolution pixel from which they were derived. For each block of low-resolution pixels, an average score $A(x)$ is used to measure how well the restored high-resolution pixels meet the average constraint imposed by their corresponding low-resolution pixels. Fig. 5 shows four low-resolution pixels whose values are μ_1, μ_2, μ_3, and μ_4. The $q \times q$ group of high-resolution pixels that are being restored from pixel μ_1 are represented by $\{x_{r,c}^{(1)}, \quad 1 \le (r,c) \le q\}$. The average score for this 2×2 block is expressed by

$$A(x) = \sum_{i=1}^{4} [\mu_i - \frac{1}{q^2} \sum_{r=1}^{q} \sum_{c=1}^{q} x_{r,c}^{(i)}]^2 \tag{9}$$

where i is the index for the low-resolution pixels, μ_i is the value of each low-resolution pixel, and $x_{r,c}^{(i)}$ are the restored high-resolution pixels corresponding to pixel μ_i. The initial high-resolution image formed by using pixel replication always has an average

score of zero because it satisfies the constraint.

The first partial derivative for the group of high-resolution pixels corresponding to

pixel μ_i is equal to

$$\frac{\partial A(x)}{\partial x_{r,c}^{(i)}} = \frac{2}{q^2}[\frac{1}{q^2}\sum_{r=1}^{q}\sum_{c=1}^{q}x_{r,c}^{(i)} - \mu_i] \tag{10}$$

The second partial derivatives are simply,

$$\frac{\partial^2 A(x)}{\partial x_{r,c}^{(i)}\partial x_{r_1,c_1}^{(i)}} = \frac{2}{q^4} \qquad \forall(r,c,r_1,c_1) \in x^{(i)} \tag{11}$$

Both the first and second partial derivatives of the average score $A(x)$ are non-zero only

with respect to pixels within the $q \times q$ group of high-resolution pixels corresponding to

the low-resolution pixel from which they were expanded.

IV. SOLVING FOR THE RESTORED IMAGE

The goal of the restoration algorithm is to solve for the image block that minimizes

the scoring function $BSA(x)$ introduced in Eq. 2. Throughout this paper, a block x is

defined both as a group of 4×4 low-resolution pixels and as the $4q \times 4q$ high-resolution

pixels that are derived from them. The 4×4 size was specifically chosen because it

contains enough pixels to adequately measure text characteristics but is not too large

to be computationally burdensome. The goal of resolution enhancement is to create a

restored image with improved resolution. This is illustrated in Fig. 6 where a 4×4

block of low-resolution pixels is expanded by a factor of $q = 4$ to create a 16×16 block

of high-resolution pixels.

Pixel replication, where every value within a $q \times q$ neighborhood is identical to the corresponding low-resolution pixel, is used for the initial expansion. Each $4q \times 4q$ block of high-resolution pixels is restored independently using iterative optimization techniques described in this section to solve for the block which minimizes the BSA score. At each iteration, the first and second partial derivatives of the BSA scoring function are used to determine the image update. To avoid block boundary discontinuities only the center $3q \times 3q$ pixels, which are highlighted in Fig. 6, are updated. The entire image is therefore divided into blocks that overlap by one quarter, or $q \times 4q$ pixels, and can be restored independently. This iterative minimization of the BSA score continues until convergence is reached resulting in the restored image.

Initially, each $4q \times 4q$ block of pixels x is converted to a $(4q)^2$-long vector $\vec{x}$ using raster scanning,

$$\vec{x}[q(r-1) + c] = x(r,c) \qquad \text{for} \qquad 1 \le r, c \le q \qquad (12)$$

A small distance away from $\vec{x}$ the BSA function can be represented by its second order Taylor series approximation [21],

$$BSA(\vec{x} + \vec{\delta}) \approx BSA(\vec{x}) + [\nabla BSA(\vec{x})]\vec{\delta} + \frac{1}{2}\vec{\delta}^T H\vec{\delta} \qquad (13)$$

and the change in BSA is given by

$$\triangle BSA = [\nabla BSA(\vec{x})]\vec{\delta} + \frac{1}{2}\vec{\delta}^T H\vec{\delta} \qquad (14)$$

where $\vec{\delta}$ is the small change to the image vector $\vec{x}$, $\nabla BSA(\vec{x})$ is the gradient, and H is the Hessian matrix. The $(4q)^2 \times (4q)^2$ Hessian given below by Eq. (15) is the symmetric

matrix of mixed partial second derivatives, which shows how a change in two variables affects the BSA function.

$$H = \begin{bmatrix} \frac{\partial^2 BSA}{\partial x_1^2} & \frac{\partial^2 BSA}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 BSA}{\partial x_1 \partial x_{(4q)^2}} \\ \frac{\partial^2 BSA}{\partial x_2 \partial x_1} & \frac{\partial^2 BSA}{\partial x_2^2} & \cdots & \frac{\partial^2 BSA}{\partial x_2 \partial x_{(4q)^2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 BSA}{\partial x_{(4q)^2} \partial x_1} & \frac{\partial^2 BSA}{\partial x_{(4q)^2} \partial x_2} & \cdots & \frac{\partial^2 BSA}{\partial x_{(4q)^2}^2} \end{bmatrix} \qquad (15)$$

Since the Hessian matrix is symmetric,

$$\frac{\partial^2 BSA}{\partial x_i \partial x_j} = \frac{\partial^2 BSA}{\partial x_j \partial x_i} \qquad (16)$$

only half of the matrix needs to be computed. To maximize the function $BSA(\vec{x})$, the variables in the Hessian matrix are first made independent. To do this the Hessian is diagonalized using a similarity transform. Each eigenvector of the Hessian matrix is placed in a separate column to form a unitary eigenmatrix E. That is, the product of the eigenmatrix with its transpose is equal to the identity matrix $EE^T = I$. When the Hessian matrix is pre-multiplied by the transposed eigenmatrix and post-multiplied by the eigenmatrix, the resulting matrix $E^T H E$ is diagonal. Because the Hessian is real and symmetric, it is always diagonalizable. As an example, shown in Fig. 7 is the Hessian matrix for an 8 × 8 block of pixels. Because of the neighborhood dependence of the scoring functions, their contributions to the Hessian are near the main diagonal. The similarity transform results in the diagonalized Hessian, $E^T H E$, which is shown at right.

The Taylor series approximation to the change in the scoring function $\triangle BSA$ can now

be expressed in terms of the $(4q)^2 \times (4q)^2$ Hessian matrix H, its $(4q)^2 \times (4q)^2$ eigenmatrix E, the $1 \times (4q)^2$ gradient of the scoring function $\nabla BSA(\vec{x})$, and the $(4q)^2 \times 1$ small change in the image vector $\vec{\delta}$,

$$\triangle BSA = ([\nabla BSA(\vec{x})]E)(E^T\vec{\delta}) + \frac{1}{2}(\vec{\delta}^T E)(E^T H E)(E^T \vec{\delta}) \tag{17}$$

With the following substitutions,

$$\nabla BSA'(\vec{x}) = [\nabla BSA(\vec{x})]E \tag{18}$$

$$\vec{\delta'} = E^T \vec{\delta} \tag{19}$$

$$H' = E^T H E \tag{20}$$

Eq. (17) can be simplified to

$$\triangle BSA = [\nabla BSA'(\vec{x})][\vec{\delta'}] + \frac{1}{2}\vec{\delta'}^T H' \vec{\delta'} \tag{21}$$

The functional minimum is achieved by stepping in the direction

$$\vec{\delta'} = H'^{-1}\nabla BSA'(\vec{x}) \tag{22}$$

in the transformed domain, which is simply $\vec{\delta} = E^T \vec{\delta'}$ in the pixel domain. For each iteration, the image update $\vec{\delta'}$ is determined. The iterations continue until convergence is reached, resulting in a desired restored image.

An example of this iterative image restoration process is shown in Fig. 4. The original 4 × 4 block of pixels is expanded by a factor of $q = 4$ using pixel replication to produce a 16 × 16 high-resolution image shown in Fig. 4(a). As the iterative restoration process proceeds in Figs. 4(b-f), the image becomes more bimodal and smooth resulting in a

greatly improved image. The majority of gray pixels that occur between characters are replaced with either black or white values, resulting in a strongly bimodal distribution. The resulting image is also smooth in both the foreground and background regions while maintaing the constraint that the average of each 4 × 4 block of high-resolution pixels is close to the original value of each corresponding low-resolution pixel. The minimization procedure is completed in 30 iterations for this image~~~~~~ As the iterations proceed, the image becomes more bimodal and smooth and these two scores are reduced.~~~~~~ The average score increases as the restoration proceeds, but this score is still significantly smaller than the bimodal and smoothness scores. Minimization of the BSA score produces a restored image that is the optimal combination of these bimodal, smoothness, and average measures.

V. EXPERIMENTAL RESULTS

The proposed BSA restoration algorithm was compared to several common expansion methods, including pixel replication, linear interpolation, and cubic spline expansion. In linear interpolation, a linear fit is calculated between all pixels within each column, and then repeated for all pixels within each row. These images naturally tend to be smooth, without sharp discontinuities, producing blurry results. Cubic spline expansion [12]

approximates the given discrete low-resolution pixels as a smooth continuous curve obtained from the weighted sum of cubic spline basis functions and resamples the curve to obtain the high resolution image. This method allows for sharp edges but often overshoots at these discontinuities, producing a ringing effect. The BSA text restoration technique creates smooth foreground and background regions and permits sharp edges at transition regions, while maintaining the low-resolution average constraint. Images restored with this technique are shown to be both qualitatively and quantitatively superior to other common resolution expansion methods. Two experiments were conducted to numerically compare the restoration methods. The first experiment involved creating low-resolution images from high-resolution originals, expanding the low-resolution imagery, and then measuring the distance to the originals. The second experiment involved scanning low-resolution document images, expanding the images with the various techniques, and using OCR accuracy to measure the success of the restoration.

To qualitatively illustrate the differences between these resolution expansion techniques, Fig. 5 shows resulting images obtained from linear interpolation, cubic spline expansion, and the proposed BSA-based restoration technique. The word "applications" from an image scanned at 100 dpi using 8-bit grayscale quantization is shown in Fig. 5(a) where significant blockiness is apparent. Linear interpolation by a factor of four was used to create the image in Fig. 5(b) which is very blurry and lacks good contrast. Fig. 5(c) depicts the resulting image from cubic spline expansion which has better contrast but is still not sharp at the edges. The image obtained using BSA restoration in

Fig. 6(d) has excellent contrast and sharp edges and is superior to the images obtained using other interpolation methods for this example.

These results clearly demonstrate several advantages of this technique that was designed specifically for text images. The BSA-restored image in Fig. 11(e) is strongly bimodal and has both smooth foreground and background regions.

As a final experiment, we used OCR accuracy to numerically compare our algorithm against existing resolution expansion methods. A set of 122 full-page journal documents from the University of Washington English Document Image Database I CD-ROM were

Fig. 10(d) has excellent contrast and sharp edges and is superior to the images obtained using other interpolation methods for this example.

The first experiment to quantitatively measure image restoration success involved creating low-resolution images by block-averaging images as described by Eq. (1). Restored images are then compared with the original to determine the success of restoration numerically. The mean squared error (MSE) was used to compare the various methods of image resolution expansion. The definition of mean squared error ▬▬▬▬▬▬ is

$$MSE = \frac{1}{RC} \sum_{r=1}^{R} \sum_{c=1}^{C} (original_{r,c} - restored_{r,c})^2 \qquad (23)$$

where R and C are the number of rows and columns in the images.

These results clearly demonstrate several advantages of this technique that was designed specifically for text images. The BSA-restored image in Fig. 11(e) is strongly

used. These binary documents were initially printed by a laser printer with 300 dots per inch (dpi) resolution. Each page was then scanned using 8-bit grayscale quantization at 75 dpi to create low-resolution original images. These 75 dpi resolution pages were then expanded using various resolution expansion methods by a factor of four to create 300 dpi images which were processed by Caere's OmniPage Pro 7.0 commercial OCR package, the world's best-selling desktop OCR software. For all cases within these experiments, both spline interpolation and BSA restoration produced superior results to linear interpolation, therefore the linear interpolation results will not be discussed further. The resulting text files were compared to the ground truth provided by the University of Washington CD-ROM using the OCR Accuracy Report Version 5.3 software developed at UNLV-ISRI. There were a total of 339,575 characters in these 122 images. Cubic spline interpolation resulted in 36,959 character errors and the BSA-based method had 30,668 character errors for an overall improvement of 17.0% for this set of images.

OCR character accuracy for all 122 document images is plotted in Fig. 8 for cubic spline interpolation and BSA expansion. The results are sorted based on the spline OCR results which are shown as a solid line. For each image, the BSA result is plotted as an "X" if the OCR accuracy is worse than spline and plotted as an "O" where the accuracy has been improved. The BSA restoration resulted in improved OCR accuracy for 72% of the images in this test set. Even in the cases where cubic spline resolution expansion improved OCR more than the BSA algorithm, the images produced by the BSA were typically more visually appealing. For the images in this experiment, the expansions produced by the BSA algorithm produce superior OCR accuracy results compared to other existing methods.

Binary Image Resolution Expansion

Document images have traditionally been scanned using binary thresholding due to the inherent binary nature of text images. Binary image processing therefore still retains great significance in the document research community. The BSA restoration method described in the previous section is well suited to process grayscale images but is not capable of restoring binary images. In order for our system to restore binary text images, each original binary image is first convolved with a spatial mask in the PreProcessing Module shown in Fig. 2 (201) to create a grayscale image. The Bimodal Estimation, Resolution Expansion, and BSA Restoration Modules used for binary images are identical to those used for grayscale images in our system. Once the high-resolution grayscale image has been created, a global threshold T_{binary} is used in the PostProcessing Module to convert the grayscale image back to a binary image. This threshold is computed to

be halfway between the white and black bimodal peaks

$$T_{binary} = \frac{\mu_{white} + \mu_{black}}{2} \tag{9}$$

This section will detail the conversion from a binary image to a grayscale image using a spatial convolution mask in the PreProcessing Module.

4.1   Binary to Grayscale Conversion

The first step of the binary restoration process is to convert the original low-resolution binary image into a low-resolution grayscale image. In our system, 8-bit grayscale quan-

tization is used so that the allowable range of values for each grayscale pixel are integers from 0 (black) to 255 (white). Convolution is performed to create gray values from a group of neighboring binary pixels. The discrete convolution of two images $f_{x,y}$ and $g_{x,y}$, denoted by $f_{x,y} * g_{x,y}$, is defined by

$$f_{x,y} * g_{x,y} = \sum_{r=0}^{R-1} \sum_{c=0}^{C-1} f_{r,c} g_{x-r,y-c} \qquad (10)$$

where R is the number of rows and C is the number of columns within each image.

The expression in Eq. (10) is a mathematical representation of a spatial convolution mask. Because convolution involves flipping $g_{x,y}$, it must be symmetric about its center. Throughout this paper the 3×3 mask $g_{x,y}$ defined by

$$g_{x,y} = \frac{1}{w_M + 4w_S + 4w_C} \begin{bmatrix} w_C & w_S & w_C \\ w_S & w_M & w_S \\ w_C & w_S & w_C \end{bmatrix} \qquad (11)$$

is used to perform this convolution, where w_M is the relative weight for the middle pixel, w_S represents the four side pixels, and w_C corresponds to the weight for each of the four corner pixels. The values of the mask are equal to zero for all other possible locations. Spatial convolution can be thought of as a weighted averaging over a neighborhood of pixels. A constraint is enforced on these weights, based on the distance from the center

of the mask,

$$w_M \geq w_S \geq w_C \geq 0 \qquad (12)$$

so pixels closer to the center are more heavily weighted. The value of the grayscale pixels x^{gray} are easily computed from the original binary pixel $x_{r,c}$ along with its eight neighbors

$$x_{r,c}^{gray} = \frac{1}{w_M + 4w_S + 4w_C}[w_M x_{r,c} + w_S(x_{r-1,c} + x_{r+1,c} + x_{r,c-1} + x_{r,c+1}) + \qquad (13)$$

$$w_C(x_{r-1,c-1} + x_{r-1,c+1} + x_{r+1,c-1} + x_{r+1,c+1}))]$$

4.2   *Constrained Binary to Grayscale Conversion*

In our system, we impose a further constraint on the computed grayscale pixels. If the binary pixel is black, then the grayscale pixel should be dark as well. Specifically, if $x = 0$ then $0 \leq x^{gray} \leq 127$ and similarly if $x = 255$ then $128 \leq x^{gray} \leq 255$.

However, this constraint is not always the case by performing the straightforward 3 × 3 convolutions as shown in Figs. 8(b-c) where the tail pixel of the character "e" which is black in Fig. 8(a) has a very light value of 198 in Fig. 8(b) and 144 in Fig. 8(c).

In order to compute each grayscale pixel while enforcing this constraint, a parameter Δ is introduced to measure the effect of the surrounding eight pixel neighbors on pixel $x(r, c)$

$$\Delta = \frac{1}{4w_S + 4w_C} \left[w_S(x_{r-1,c} + x_{r+1,c} + x_{r,c-1} + x_{r,c+1}) + \right. \tag{14}$$
$$\left. w_C(x_{r-1,c-1} + x_{r-1,c+1} + x_{r+1,c-1} + x_{r+1,c+1}) \right]$$

The range of this measure is $0 \leq \Delta \leq 255$. The computed grayscale pixel x^{gray} is then determined based on the original value of the binary image x in the following manner

$$\text{if} \quad x = 0 \quad \text{then} \quad x^{gray}(r, c) = \lfloor \frac{\Delta}{2} \rfloor \tag{15}$$
$$\text{if} \quad x = 255 \quad \text{then} \quad x^{gray}(r, c) = 128 + \lfloor \frac{\Delta}{2} \rfloor$$

where $\lfloor x \rfloor$ is the integer floor rounding function defined by the largest integer that is less than or equal to x.

To quantitatively compare our system with other expansion methods, resulting images were compared with cubic spline expansion to measure OCR performance. Shown in Fig. 6 is the resulting 300 dpi image obtained using cubic spline expansion on a paragraph from one of the University of Washington document images. The text file created by OCR is shown in Fig. 6(b) where mistakes have been highlighted. For this example, the OCR results for the cubic spline image had 7 areas where mistakes were made. The same image paragraph was processed by our system resulting in the image in Fig. 7(a) which has better contrast than its corresponding cubic spline image in Fig. 6(a). OCR results are improved as well, only 4 mistake areas are highlighted in Fig. 7(b).

PAGE 33 INTENTIONALLY OMITTED

Experiment -- University of Washington Binary Documents

To measure the success of the system in restoring binary document images, a second experiment was conducted using a total of 48 document pages from the University of Washington database. These images were scanned at 100 dpi using binary quantization and then expanded by a factor of three using the binary BSA restoration technique described in Section 4 to create 300 dpi images. The binary enhancement process for the sample word "representative" from this experiment is shown in Fig. 8 to visually illustrate the improvements. The original binary image in Fig. 8(a) was convolved with a spatial mask of weights $w_S = 1$ and $w_C = 0.707$ to produce a grayscale image in Fig. 8(b). This low-resolution grayscale image was enhanced by the system to produce a high-resolution grayscale image in Fig. 8(c) which was converted to a binary image to produce a restored image in Fig. 8(d). The resulting high-resolution image was noticably less blocky than the original.

To quantitatively measure the success of the system in enhancing binary document images, OCR accuracy was compared before and after resolution expansion. The expanded images produced by our system were compared with 300 dpi images created by pixel replication. There were a total of over 140,000 characters in this dataset. The overall

character accuracy for the original images was 82.0% and the overall character accuracy for the restored images was 89.1% which was a 39.5% reduction in errors, a significant improvement over pixel replication. (See Fig 8 (b))

5.3   System Performance

The previous experimental results demonstrate the advantages of our text restoration system compared to standard methods of interpolation such as linear interpolation and cubic spline expansion. The system was tested using images from the widely avaliable University of Washington document database and evaluated using the commercial Caere OmniPage Pro OCR software package to measure character accuracy. For the document and video images in these experiments, the images produced by our restoration system resulted in a higher OCR accuracy compared to other standard methods.

As stated earlier ████████, the goal of this resolution expansion system is to improve the OCR accuracy. Real-time image processing is not a requirement. Our system takes approximately 10 minutes to expand a single full-page 100 dpi document image by a factor of 3 on a 250 MHz workstation. Cubic spline interpolation is much faster than our system and takes about 10 seconds for a single document page. Video frames are typically smaller than low-resolution document images and are therefore processed more quickly by the system. A 320 × 240 video frame takes approximately 2 minutes for our system to expand by a factor of three. Cubic spline expansion of a video frame only takes about 2 seconds. If the speed concern becomes an issue in the future, our approach has a potential for massive parallelization because the images can

be divided into blocks of pixels which can be restored independently.

Thus, Our system was shown to be capable of enhancing grayscale documents and video images as well as binary document images. The success of the system was demonstrated by experiments using images from a standard document image database and a commercial OCR package. Restoration of grayscale images was performed by optimizing bimodal, smoothness, and average (BSA) scores that measure desired properties of text images. These scores were combined to form a single scoring function which produced images that were strongly bimodal and smooth, while satisfying the average constraint score. When the original image was binary, it was initially converted to a grayscale image using a spatial convolution mask and then processed as grayscale images. These resultant images restored using our system were shown to be superior to images expanded using existing linear interpolation and cubic spline expansion techniques.

Optical character recognition accuracy was also used to quantitatively measure the success of various resolution expansion methods. Both binary and grayscale text images restored with our system were shown experimentally to improve OCR accuracy compared to linear interpolation and cubic spline expansion. Despite that our system was designed to enhance text images, it can also be used to accurately expand checks, tables, charts, graphs, line drawings, and other types of documents as well due to their similarities.

References

[1] P. Stubberud, J. Kanai, and V. Kalluri, "Adaptive image restoration of text images that contain touching or broken characters," *ICDAR95 Proceedings International Conference on Document Analysis Recognition*, pp. 778-781, 1995.

[2] A. P. Whichello and H. Yan, "Linking broken character borders with variable sized masks to improve recognition," *Pattern Recognition*, Vol. 29, No. 8, pp. 1429-1435, 1996.

[3] M. Y. Jaisimha, E. A. Riskin, R. Ladner, and S. Werner, "Model-based restoration of document images for OCR," *Proceedings of the SPIE*, Vol. 2660, pp. 297-308, 1996.

[4] M.Y. Yoon, S.W. Lee, and J.S. Kim, "Faxed image restoration using Kalman filtering," *ICDAR95 Proceedings International Conference on Document Analysis Recognition*, pp. 677-681, 1995.

[5] J. Liang, R. M. Haralick, and I. T. Phillips, "Document image restoration using binary morphological filters," *Proceedings of the SPIE*, Vol. 2660, pp. 274-285, 1996.

[6] T. Akiyama, N. Miyamoto, M. Oguro, and K. Ogura, "Faxed document image restoration method based on local pixel patterns," *Proceedings of the SPIE*, Vol. 3305, pp. 253-262, 1998.

[7] G. Ramponi and P. Fontanot, "Enhancing document images with a quadratic filter," *Signal Processing*, Vol. 33, pp. 23-34, 1993.

[8] L. Koskinen, H. Huttunen, and J. T. Astola, "Text enhancement method based on soft morphological filters," *Proceedings of the SPIE*, Vol. 2181, pp. 243-253, 1994.

[9] Y.C. Shin, R. Sridhar, V. Demjanenko, P. W. Palumbo, and J. J. Hull, "Contrast enhancement of mail piece images," *Proceedings of the SPIE*, Vol. 1661, pp. 27-37, 1992.

[10] F. Sattar and D. B. H. Tay, "On the multiresolution enhancement of document images using fuzzy logic approach," *ICPR98 Proceedings International Conference on Pattern Recognition*, pp. 939-942, 1998.

[11] T. C. Chen and R. J. P. de Figueiredo, "Image decimation and interpolation techniques based on frequency domain analysis," *IEEE Transactions on Communications*, Vol. 32, No. 4, 1984.

[12] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 26, No. 6, December 1978.

[13] N. B. Karayiannis and A. N. Venetsanopoulos, "Image interpolation based on variational principles," *Signal Processing*, Vol. 25, pp. 259-288, 1991.

[14] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 29, No. 6, pp. 1153-1160, 1981.

[15] A. D. Kulkarni and K. Sivaraman, "Interpolation of digital imagery using hyperspace approximation," *Signal Processing*, Vol. 7, pp. 65-73, 1987.

[16] V. S. Nalwa, "Edge-detector resolution improvement by image interpolation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 3, pp. 446-451, 1987.

[17] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, Vol. 3, No. 3, pp. 233-242, 1994.

[18] H. Li, O. E. Kia, and D. S. Doermann, "Text enhancement in digital video," *Proceedings of the SPIE*, Vol. 3651, pp. 2-9, 1999.

[19] M. J. Taylor and C. R. Dance, "Enhancement of document images from cameras," *Proceedings of the SPIE*, Vol. 3305, pp. 230-241, 1998.

[20] R. W. Schafer, R. M. Mersereau, and M. A. Richards, "Constrained iterative restoration algorithms," *Proceedings of the IEEE*, Vol. 69, No. 4, April 1981, pp. 432-450.

[21] J. Skilling and R. K. Bryan, "Maximum entropy image reconstruction: general algorithm," *Mon. Not. Royal Astronomy Society*, Vol. 211, 1984, pp. 111-124.

[22] P. D. Thouin and C. -I Chang, "A Method for Restoration of Low-Resolution Text Images," *Symposium on Document Image Understanding*, pp. 143-148, 1999.

References

[1] Caere Corporation, "OmniPage Pro 9.0 Software Package," (1998).

[2] Xerox Corporation, "TextBridge Pro 9.0 Software Package," (1999).

[3] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29(6), pp. 1153-1160 (1981).

[4] T. C. Chen and R. J. P. de Figueiredo, "Image decimation and interpolation techniques based on frequency domain analysis," *IEEE Transactions on Communications* 32(4), pp. 479-484 (1984).

[5] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(1), pp. 39-55 (1987).

[6] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6(6), pp. 721-741 (1984).

[7] A. D. Kulkarni and K. Sivaraman, "Interpolation of digital imagery using hyperspace approximation," *Signal Processing* 7, pp. 65-73 (1987).

[8] V. S. Nalwa, "Edge-detector resolution improvement by image interpolation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(3), pp. 446-451 (1987).

[9] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26(6), pp. 508-517 (1978).

[10] N. B. Karayiannis and A. N. Venetsanopoulos, "Image interpolation based on variational principles," *Signal Processing* 25, pp. 259-288 (1991).

[11] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing* 3(3), pp. 233-242 (1994).

[12] M. Shridhar, M. Ahmadi, and M. El-Gabali, "Restoration of noisy images modeled by Markov random fields with Gibbs distributions," *IEEE Transactions on Circuits and Systems* 36(6), pp. 884-890 (1989).

[13] P. D. Thouin and C. -I Chang, "A Method for Restoration of Low-Resolution Text Images," *Symposium on Document Image Understanding*, pp. 143-148 (1999).

[14] J. Skilling and R. K. Bryan, "Maximum entropy image reconstruction: general algorithm," *Mon. Not. Royal Astronomy Society* 211, pp. 111-124 (1984).

[15] R. W. Schafer, R. M. Mersereau, and M. A. Richards, "Constrained iterative restoration algorithms," *Proceedings of the IEEE* 69(4), pp. 432-450 (1981).

US Provisional Application Serial No. 60/127,316, filed April 1,
1999 (including its attachments) and the following attached
documents are incorporated herein by reference as if fully set
forth herein:

1. "A Method for Restoration of Low-Resolution Bimodal Images"
(authored by CHEIN-I CHANG and PAUL THOUIN, unpublished, but
scheduled for presentation to 1999 Symposium on Documentation
Image Understanding Technology at Annapolis, MD April 14-16,
1999)

2. "Constrained nonlinear restoration of JPEG compressed low-
resolution text from grayscale images using a Gibbs-Markov random
field prior", authored by CHEIN-I CHANG and PAUL THOUIN, and
published in IS&T/SPIE 10th Annual Symposium, Electronic Imaging
'98: Science & Technology (presented at San Jose, California
January 24-30, 1998); and

3. "A Gibbs-Markov Random Field Approach to Restoration of JPEG-
compressed Text Images" (unpublished, authored by CHEIN-I CHANG
and PAUL THOUIN).

While a specific embodiment of the invention has been shown and
described in detail to illustrate the application of the
principles of the invention, it will be understood that the
invention may be embodied otherwise without departing from such

43

principles and that various modifications, alternate constructions, and equivalents will occur to those skilled in the art given the benefit of this disclosure. Thus, the invention is not limited to the specific embodiment described herein, but is defined by the appended claims.

CLAIMS

We claim:

1. A process for restoration of a low resolution image comprising the steps of:

    subdividing said low-resolution image into groups of pixels;

    generating initial conditions as an initial expansion of a restored (high-resolution) image;

    iteratively revising each group of the restored (high-resolution) image so as to minimize the BSA score for said group of pixels.

2. The process of claim 1 wherein the step of generating initial conditions comprises the step of either replicating or linearly interpolating or cubic-spline interpolating each group of pixels.

3. The process of claim 1 wherein said low-resolution image is a binary image.

4. The process of claim 1 wherein said low-resolution image is a gray-scale image.

5. The process of claim 1 wherein said low-resolution image is a color image.

6. The process of claim 1 wherein said low-resolution image is a video image.

7. The process of claim 1 wherein an image is provided, an initial BSA score associated with said input is computed, an updated image output is produced from said image and its associated BSA score, an updated BSA score is computed associated with the updated image output, the prior computed BSA score is compared with the updated BSA score and updated image outputs and calculate updated BSA scores are produced iteratively until an updated BSA score does not differ from its prior computed BSA score by more than a predetermined amount.

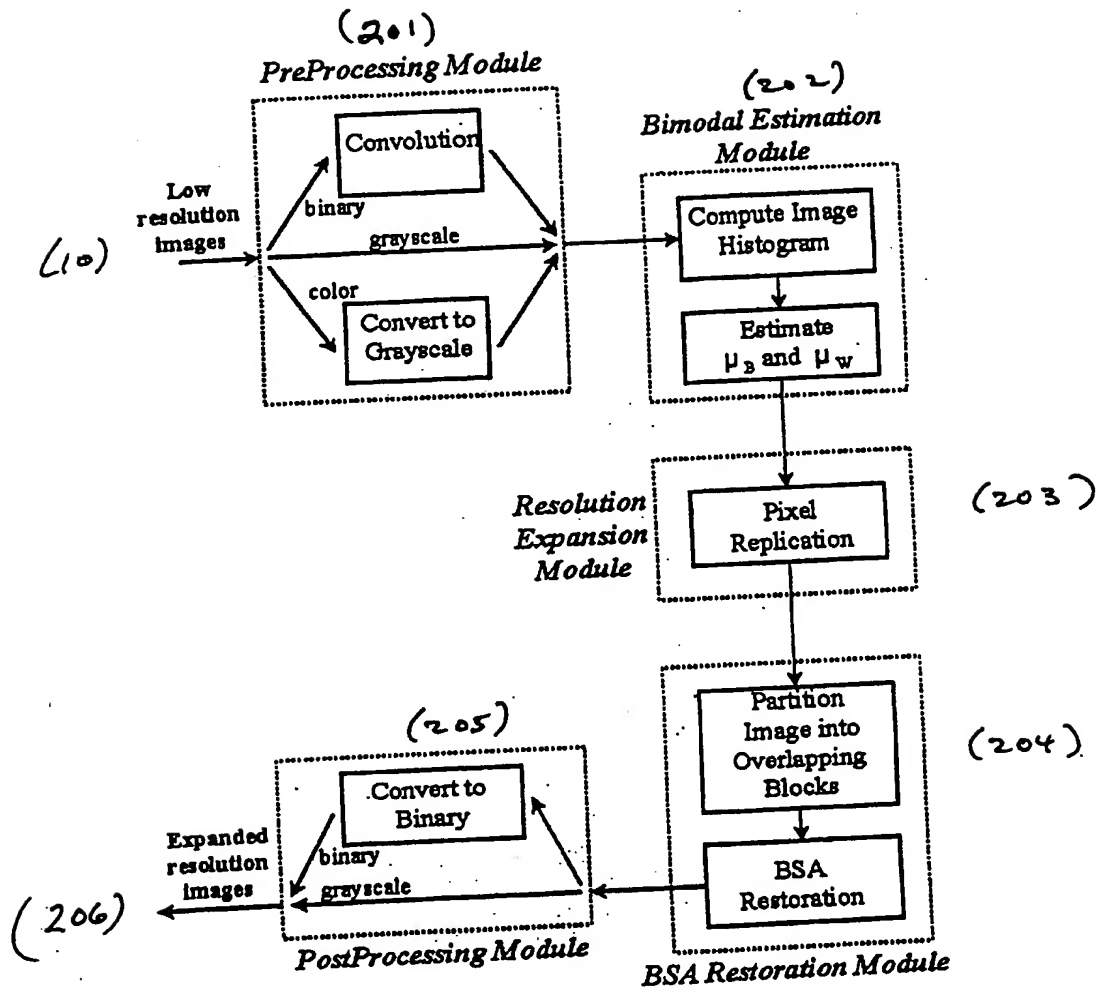8. A machine for restoration of a low resolution image comprising image acquisition means for providing said low resolution image to a computer processor, said processor programmed to carry out the steps of subdividing said low-resolution image into groups of pixels; replicating each group of pixels as an initial expansion of a restored (high-resolution) image; iteratively revising each group of the restored (high-resolution) image so as to minimize the BSA score for said group of pixels; and providing as an output a restored image.

9. A machine as in claim 8 wherein said computer processor further includes a preprocessing module which provides an output to a bimodal estimation module which outputs an estimation to a

46

resolution expansion module for pixel replication and output to a
BSA restoration module which produces an output representative of
an enhanced image.

10.   A machine as in claim 8 wherein said BSA restoration module
comprises means for accepting an image as an input, computing an
initial BSA score associated with said input and producing an
updated image output therefrom, computing an updated BSA score
associated with the updated image, comparing the prior computed
BSA score with the updated BSA score and continuing to produce
updated image outputs and calculate updated BSA scores until an
updated BSA score does not differ from its prior computed BSA
score by more than a predetermined amount.

CHANG / THOUIN

FIG 1

IMAGE

Sensors (2)

Image Acquisition System
(1)

Image Data (10)

Input Means (101)

General Purpose Computer (100)

Processing Means
(102)

Output Means (103)

Output (enhanced image)

(201)
PreProcessing Module

(202)
Bimodal Estimation Module

Low
resolution
images

(10)

binary

grayscale

color

Convolution

Convert to
Grayscale

Compute Image
Histogram

Estimate
μ_B and μ_W

*Resolution
Expansion
Module*

Pixel
Replication

(203)

(205)

Convert to
Binary

Expanded
resolution
images

(206)

binary

grayscale

PostProcessing Module

Partition
Image into
Overlapping
Blocks

(204)

BSA
Restoration

BSA Restoration Module

Figure 2. Block diagram of the text resolution expansion system.
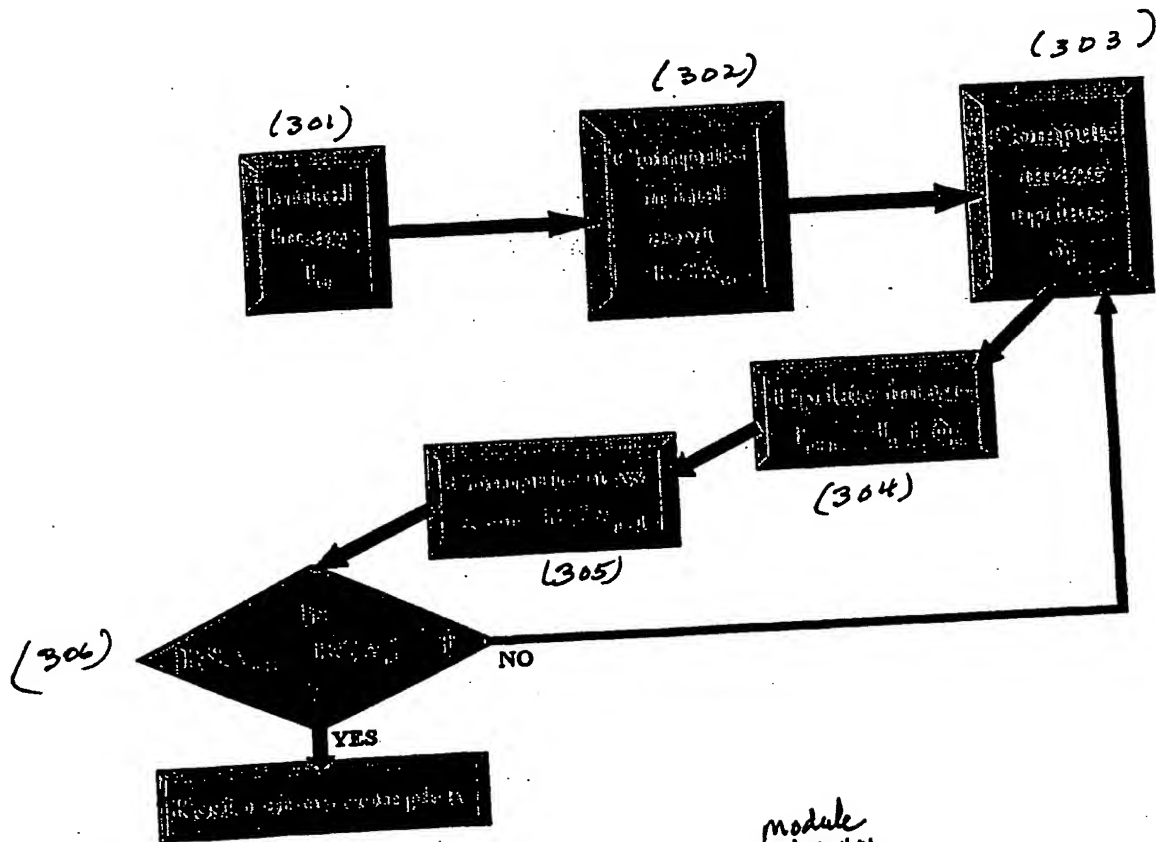
Figure 3: BSA restoration module

(a) Original   (b) Iteration 3   (c) Iteration 6

(d) Iteration 10   (e) Iteration 15   (f) Iteration 30
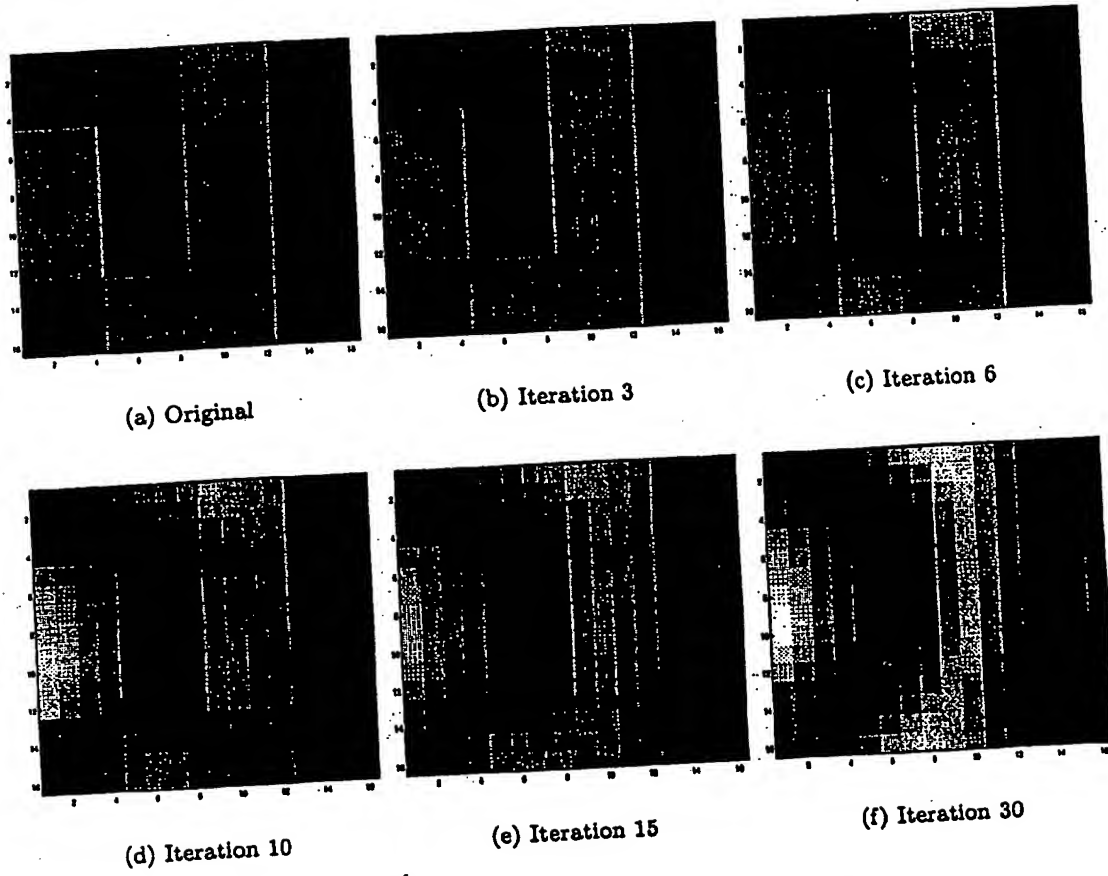
Figure 6: Iterative BSA restoration example.

(a) Original image
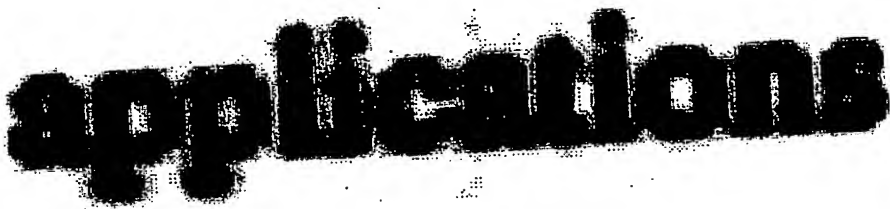
(b) Linear interpolated image

(c) Cubic spline image
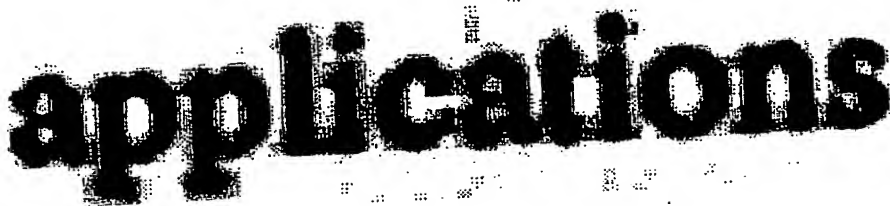
(d) BSA restored image

Figure 6: Results of grayscale document restoration.

(a) Cubic spline image

The Preisach Model for ferromag.-Ats is generalized and adapted for the description of the hysteretic behaviour of a polycrystallint specinm of shape memory alloys. The thanwdynamical properties of the individual crystallites are described by the Landau-Devonshire free energy which contains four parameters. The corresponding quadruplets of parameters of a polycrystalline body fill a region in a four-dimensional Preisach space. A thermodynamical loading path will sweep surfaces across this region and change phases in the process. The physical problem of the response of a specimen to applied loads is thus convened into the geometrical problem of counting volumes between moving surfaces. This conversion facilitates the numerical evaluation of the eftet of complicated loading pathL

(b) Cubic spline OCR results

Figure 🖋 Example OCR results for cubic spline.

The Preisach Model for ferromagnets is generalized and adapted for the description of the hysteretic behaviour of a polycrystalline specimen of shape memory alloys. The thermodynamical properties of the individual crystallites are described by the Landau-Devonshire free energy which contains four parameters. The corresponding quadruplets of parameters of a polycrystalline body fill a region in a four-dimensional Preisach space. A thermodynamical loading path will sweep surfaces across this region and change phases in the process. The physical problem of the response of a specimen to applied loads is thus converted into the geometrical problem of counting volumes between moving surfaces. This conversion facilitates the numerical evaluation of the effect of complicated loading paths.

(a) BSA restored image

The Preiiach Model for fierromag.iets is generalized and adapted for the description of the hysteretic behaviour of a polycrystalline specimen of shape memory alloys. The thermodynamical properties of the individual crystallites are described by the Landau-Devonshire free energy which contains four parameters. The corresponding quadruplets of parameters of a Polycrystalline body fill a region in a four-dimensional Preisach space. A thermodynamical loading path will sweep surfaces across this region and change phases in the process. The physical problem of the response of a specimen to applied loads is thus converted into the geometrical problem of counting volumes between moving surfaces. This conversion facilitates the numerical evaluation of the effect of complicated loading paths.

(b) BSA restored OCR results
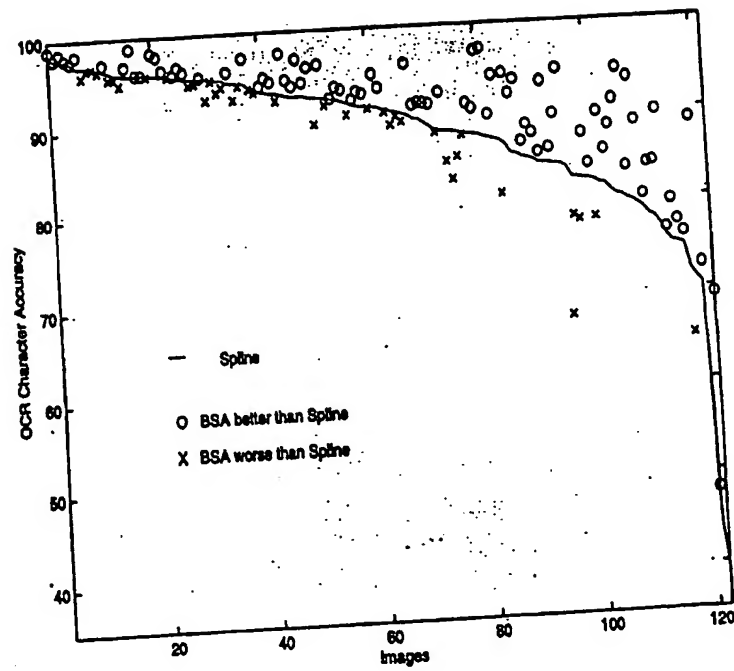
Figure 7: Example OCR results for BSA.

Figure 8: OCR accuracy - spline vs. BSA

representative

(a) Original binary image
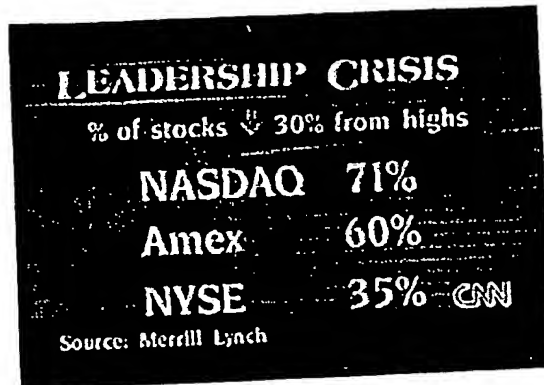
representative

(b) Convolved grayscale image

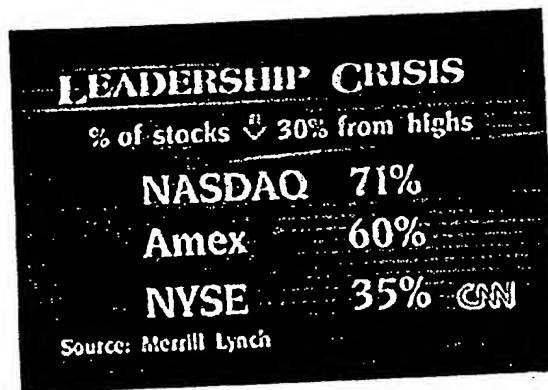representative

(c) Restored grayscale image
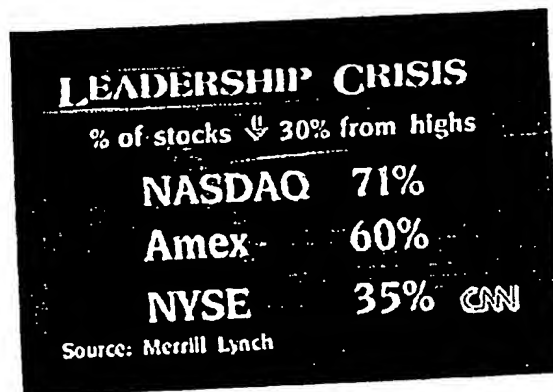
representative

(d) Restored binary image

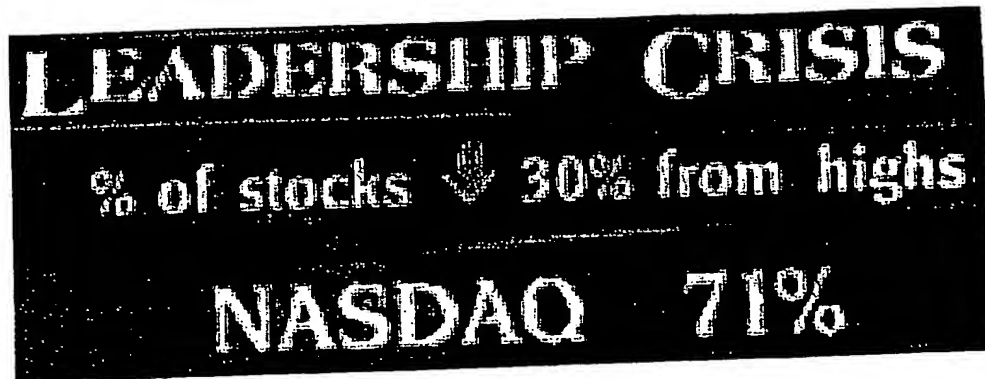Figure 9: Results of binary document restoration.
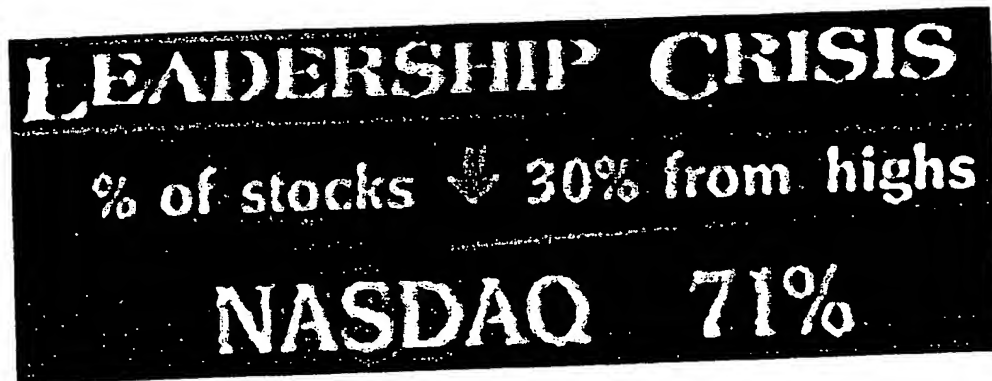
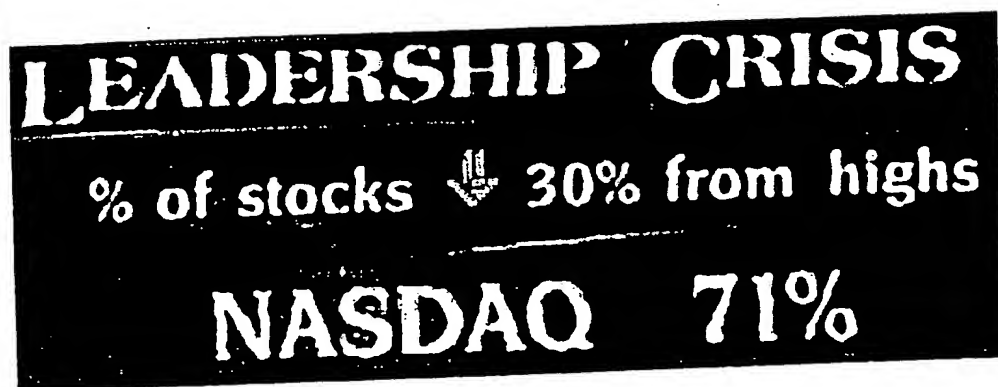(a) Original image

(b) Cubic spline image

(c) BSA-restored image

Figure 10: Examples of video frame restoration results

(a) Original image

(b) Cubic spline image

(c) BSA-restored image

Figure 8: Part of enlarged video frame restoration result

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :GO6K 9/32
US CL   :382/299

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 382/298-301; 358/1.2, 451, 528

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST

search terms: bimod$, smooth$, BSA, resolut$, pixel$, scal$, magnif$, reduc$

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,611,023 A (HANYU) 11 March 1997, col. 3, lines 44-64.	1-10
A	US 5,650,858 A (LUND) 22 July 1997, Abstract.	1-10
A	US 5,586,196 A (SUSSMAN) 17 December 1996, col. 4, lines 19-37.	1-10
A,P	US 5,991,448 A (GO) 23 November 1999, col. 8, lines 28-56.	1-10

☐ Further documents are listed in the continuation of Box C.   ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 JUNE 2000	**0 6 JUL 2000**

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Stephen Brinich
Facsimile No. (703) 305-3230	Telephone No. (703) 305-4390

Form PCT/ISA/210 (second sheet) (July 1998) ★